AD-A099 291 STANFORD UNIV CA SYSTEMS OPTIMIZATION LAB F/G 12/1
CONVERGENCE RATES OF THE ELLIPSOID METHOD ON GENERAL CONVEX FUN--ETC(U)
JAN 81 J GOFFIN N00014-75-C-0267
UNCLASSIFIED SOL-81-3 NL

1 OF 1
AD
A099291

END
DATE
FILMED
8-81
DTIC

LEVEL II

Systems
Optimization
Laboratory

AD A099291

DTIC FILE COPY

This document has been approved for public release and sale; its distribution is unlimited.

DTIC
ELECTE
MAY 26 1981
S D
A

Department of Operations Research
Stanford University
Stanford, CA 94305

81 5 18 121

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

CONVERGENCE RATES OF THE ELLIPSOID METHOD ON GENERAL CONVEX FUNCTIONS

by

Jean-Louis Goffin*

TECHNICAL REPORT SOL 81-3

January 1981

*McGill University, Montreal, Canada. This research was done while the author was visiting the Systems Optimization Laboratory, Department of Operations Reseach at Stanford Univeristy.

Research and reproduction of this report partially were supported by the Department of Energy Contract AM03-76SF00326, PA# DE-AT03-76ER72018, Office of Naval Research Contract N00014-75-C-0267, National Science Foundation Grants MCS76-81259, MCS-7926009, and ECS-8012974 at Stanford University; and by the DGES of Quebec, the NSERC of Canada (Grant A-4152), and the SSHRC of Canada.

Abstract

The ellipsoid method is applied to the unconstrained minimization of a general convex function. The method converges at a geometric rate, which depends only upon the dimension of the space but not on the actual function. This rate can be improved somewhat if the function satisfies some Lipschitz-type condition, or if the minimum set has dimension greater than zero.

If the ellipsoid entirely contains the optimal set, equating the Steiner polynomial associated to the optimal set, and the volume of the ellipsoid at a given iteration, will give an upper bound on the minimum recorded function value.

Keywords: Ellipsoid method, nondifferentiable optimization, convex programming, volumes.

1. Introduction

The ellipsoid method is an implementable version of the method of centers of gravity of Levin [13]; both algorithms are designed to minimize a general convex function.

Levin's method is a natural extension to general convex functions, in n dimensions, of the bisection algorithm for one dimensional unimodal functions. It can also be viewed as a cutting plane method, where the new iterate is defined as the center of gravity of the set formed by the intersection of all previously generated cutting planes; a new cutting plane is generated at this point, and added to the list of cutting planes. What is quite remarkable about this is that the domain of localization of the solution (i.e., the intersection of the previously generated hyper-planes) has its volume reduced at each iteration by a ratio of at least $1 - (n/(n+1))^n < 1 - 1/e$. This implies that the minimum recorded function value is bounded above by a geometric progression of ratio $(1 - 1/e)^{1/n}$; this result is not quite proved in Levin's paper.

The ellipsoid method is also a particular version of a class of algorithms due to Shor [17, 18, 19, 20, 21] which can be described as a variable metric subgradient optimization method, where the metric is updated at every iteration by a rank one matrix. The variable metric was introduced by Shor as an attempt to correct the bad convergence of subgradient optimization on convex functions with very elongated level sets (or badly conditioned, acute, very kinky, gully-type functions). Clearly any variable metric method can be interpreted in terms of an ellipsoid. The method worked reasonably well as a heuristic, but proofs of convergence were hard to come by, and quite unsatisfactory.

Yudin and Nemirovskii, in two seminal papers [22, 23] dealing with the computational complexity of the general convex programming problem, did (among other things) combine the methods of Shor and Levin in what is now called the ellipsoid method: it is an implementable version of Levin's method, and also a version of Shor's method for which a proof of convergence exists. In restricting oneself to ellipsoids rather than using general convex sets, the volume of the domain of localization of the solution is reduced at each iteration, by the ratio

$$\frac{n}{n+1}\left(\frac{n^2}{n^2-1}\right)^{(n-1)/2} < e^{-[1/2(n+1)]} ;$$

this implies that the ellipsoid method would take approximately $2n/e$ times the number of iterations that Levin's method requires to reach a given accuracy. Yudin and Nemirovskii also show that no algorithm (using the information given by an oracle which returns the value of the function and a subgradient corresponding to each iterate) can improve significantly on Levin's method.

The results given by Yudin and Nemirovskii are of a slightly different nature from the ones given in this paper; they look for a minimum of a convex function subject to the restrictions that the point belongs to a given compact convex set G (with an interior) and that if satisfies one convex inequality (or any finite number of convex inequalities). The presence of one inequality has the impact that the algorithm depends upon the precision that one decides to reach at termination, and thus no rate of convergence can be given. The presence of the set G (required to make

a definition of computational complexity work) is handled by assuming that an explicit projection map on G is available, thus restricting G to be a simplex, a sphere, a rectangular parallelotope, etc. ..., but not an ellipsoid (the projection on an ellipsoid appears to require the solution of an eigenvalue-vector problem); the algorithm is also slightly changed.

For the problem of minimizing, without constraints, a general convex function, Shor [20] showed that convergence is bounded by the product of an arithmetic series and a geometric series.

Khacian [10, 11] applied the results of Yudin and Nemirovskii to the linear programming problem, or to the problem of solving a system of linear inequalities. Again the convergence proof is based upon a perturbation technique, and no rates of convergence are given.

In Section 2 of this paper, it is first shown that the n^{th} root of the volume of the intersection of the level sets of a convex function with a convex set is a concave function of the level. This implies, using a parametrized version of the proof given by Yudin and Nemirovskii [22, 23], Khacian [10, 11], Gacs and Lovasz [8], and Aspvall and Stone [2], that convergence is finite if the initial ellipsoid intersects the optimal set in a set of dimension n, while if this intersection is not empty convergence occurs at a geometric rate which is approximately $1 - 1/2n^2$.

This result is valid for any convex function, and is independent of the particular function. In Section 3, an attempt is made at showing that the rate of convergence of the ellipsoid method may depend on the properties of the function one minimizes: it depends both on some Lipschitz-type characteristics of the function, and on the dimensionality of the intersection

of the initial ellipsoid with the optimal set. It is also shown that a bound on the value of the function reached is, if the initial ellipsoid entirely contains the optimal set, given by an equation relating the Steiner polynomial of the optimal set to the volume of the ellipsoid, at a given iterate.

Section 4 is e. sentially an appendix containing a few necessary technical results on volumes. It is based upon the books of Bonnensen and Fenchel [3], Busemann [4], Eggleston [5], and Hadwiger [9].

2. A general convergence theory

The ellipsoid algorithm will be used to solve the problem of minimizing a general convex function f defined on R^n. Let $\partial f(x)$ be the subdifferential of f at x, $f^* = \mathrm{Inf}\{f(x) : x \in R^n\}$ be the minimum value of f, and $S^* = \{x \in R^n : f(x) \leq f^*\} = \{x \in R^n : 0 \in \partial f(x)\}$ be the set of minimal points. It will be assumed throughout (unless otherwise specified) that f^* is finite and that S^* is not empty.

Let $d(x,S)$ denote the Euclidean distance between a point x and a set S, i.e, $d(x,S) = \mathrm{Inf}\{\|x-y\| : y \in S\}$ where $\|\cdot\|$ is the Euclidean norm; $d(x)$ will be used for $d(x,S^*)$. The diameter of a compact set S is defined by $D(S) = \mathrm{Sup}\{\|x-y\| : x \in S, y \in S\}$; while the inradius $r(S)$ is the radius of a largest sphere contained in S. The unit closed ball is $B = \{x \in R^n : \|x\| \leq 1\}$; the interior of B is denoted by B^0. The set of all compact nonempty sets in R^n is metrized by the Hausdorff distance

$$d(A_1,A_2) = \mathrm{Inf}\{\delta \geq 0 : A_1 + \delta B^0 \subset A_2,\ A_2 + \delta B^0 \subset A_1\} \quad ;$$

convergence and continuity are defined using this metric (or any of its topologically equivalent versions).

The volume $V[S]$ of a bounded measurable set S in R^n is its n dimensional Lebesgue measure; $V[\phi]$ will be taken as $-\infty$. The volume of the unit ball is $\omega_n = \pi^{n/2}/\Gamma(n/2 + 1)$. The surface of A is denoted by $F(A)$. If A is a s dimensional set in R^n, then $V_s[A]$ (and $F_s[A]$) will denote the volume (and the surface) of A within the s dimensional affine manifold containing A.

The ellipsoid algorithm:

The ellipsoid algorithm computes a sequence $\{(x^{(k)}, E_k) : k = 0,1,\ldots\}$ where each E_k is an ellipsoid in R^n and $x^{(k)}$ is its center.

Step 1

Choose a point $x^{(0)} \in R^n$ and an ellipsoid $E_0 \subset R^n$ centered at $x^{(0)}$ such that $E_0 \cap S^*$ is not empty (i.e., $\mathrm{Min}\{f(x) : x \in E_0\} = f^*$).

Step 2

Compute $a^{(k)} \in \partial f(x^{(k)})$.

If $a^{(k)} = 0$, stop with $x^{(k)}$ optimal; otherwise go to Step 3.

Step 3

Let $H_k = \{x \in R^n : (a^{(k)}, x-x^{(k)}) \leq 0\}$; then define E_{k+1} as the least volume ellipsoid cointaining $H_k \cap E_k$, and take $x^{(k+1)}$ as the center of E_{k+1}. Set $k \leftarrow k+1$ and go to Step 2.

This algorithm is implementable as the ellipsoids can be described by matrices, and the iteration from E_k to E_{k+1} is described by a rank one matrix update [2, 8, 10, 11, 17, 18, 19, 20, 21].

The key observation is that

$$V[E_{k+1}] = c_n V[E_k] \quad \text{and thus} \quad V[E_k] = (c_n)^k V[E_0]$$

with

$$c_n = \frac{n}{n+1} \left(\frac{n^2}{n^2-1}\right)^{(n-1)/2} < e^{-1/2(n+1)} \;;$$

and $V[E_k]$ decreases as a geometric series whose ratio depends only upon the dimension of the space.

It should be remarked that in the remainder of this paper the fact that the sets E_k are ellipsoids is not used (at least not much). It follows that most results can be applied to Levin's method of centers of gravity with $\tilde{c}_n = 1 - (n/[n+1])^n < 1 - 1/e$ replacing c_n; the method is a nonimplementable version of the ellipsoid algorithm, which uses general convex sets A_k, with inStep 3 $A_{k+1} = H_k \cap A_k$ and $x^{(k+1)}$ is defined as the center of gravity of A_{k+1} (one has $V[A_{k+1}] \leq \tilde{c}_n V[A_k]$, [14]). Levin's method is clearly a cutting plane method.

It is an open question whether other classes of convex sets (simplices, parrallelotopes, etc., ...) would provide implementable versions of this algorithm.

Proofs of convergence of the ellipsoid algorithm are based upon a study of the behavior of volumes. Other characteristics of convex sets may be used, for instance the inradius [11, 12], as clearly

$$r(E_k) \leq (V[E_k]/\omega_n)^{1/n} = (V[E_0]/\omega_n)^{1/n} (c_n)^{k/n} \quad ;$$

and furthermore $r(\{x \in E_0 : f(x) \leq \alpha\})$ is a concave function of α.

Usually the starting ellipsoid E_0 will be taken as a sphere with center $x^{(0)}$ and radius r ($E_0 = x^{(0)} + rB$); clearly in this case $E_0 \cap S^* \neq \phi$ if and only if $d(x^{(0)}) \leq r$, which means that one needs an overestimate of the distance between a point $x^{(0)}$ and the opitmal set S^*.

The general theory of convergence of the ellipsoid method will be based upon the properties of the function

$$h(\varepsilon; E_0) = V[T_{f^*+\varepsilon} \cap E_0]$$

where $T_\alpha = \{x \in R^n : f(x) \leq \alpha\}$ are the level sets of f and E_0 is the starting ellipsoid.

If the optimal set S^* is bounded then one can define the function

$$h(\varepsilon) = V[T_{f^*+\varepsilon}] \quad .$$

Clearly $h(\varepsilon) = h(\varepsilon; E_0)$ for all $\varepsilon \in [0, \varepsilon_0(E_0)]$, where

$$\begin{aligned}\varepsilon_0(E_0) &= \mathrm{Sup}\{\varepsilon : T_{f^*+\varepsilon} \subset E_0\} \\ &= \mathrm{Inf}\{f(x) - f^* : x \notin E_0\} \quad .\end{aligned}$$

Lemma 2.1:

Let f be a convex function defined on R^n, with minimum f^*, and minimum set S^*, and E be a compact convex set such that $E \cap S^*$ is not empty, then the function $h(\varepsilon; E) = V[T_{f^*+\varepsilon} \cap E]$ defined for $\varepsilon \in [0,\infty)$ has the following properties:

i) It is continuous on $\varepsilon \in [0,\infty)$.

ii) It is strictly increasing for $\varepsilon \in [0, \varepsilon^*(E)]$ where $\varepsilon^*(E) = \mathrm{Max}\{f(x)-f^*: x \in E\}$.

iii) $h(\varepsilon; E) = V[E]$ for all $\varepsilon \geq \varepsilon^*(E)$, and $h(0; E) = V[E \cap S^*]$.

iv) $h^{1/n}(\varepsilon; E)$ is concave for $\varepsilon \in [0,\infty)$.

v) $h^{1/n}(\varepsilon; E) \geq (\varepsilon/\varepsilon^*(E))\, V^{1/n}[E] + (1 - \frac{\varepsilon}{\varepsilon^*(E)})\, V^{1/n}[E \cap S^*]$ for $\varepsilon \in [0, \varepsilon^*(E)]$.

Proof:

i) follows from the continuity of $T_{f^*+\varepsilon} \cap E$ for all $\varepsilon \in [0,\infty)$ (by Theorem 4.6) and the continuity of the volume on the set of compact convex sets (by Theorem 4.8).

ii) and iii) are clear.

iv) follows from Lemma 4.4 and Theorem 4.7.

v) $h^{1/n}(\varepsilon; E) = h^{1/n}(\frac{\varepsilon}{\varepsilon^*(E)}\, \varepsilon^*(E) + (1 - \frac{\varepsilon}{\varepsilon^*(E)})0; E)$ and thus v) follows from iii) and iv). Q.E.D.

The function $h(\varepsilon)$, defined only if S^* is compact, has similar properties.

Definition 2.2:

The functions $h^{-1}(t; E)$ and $h^{-1}(t)$ are defined as the inverses of the functions $h(\varepsilon; E_0)$ and $h(\varepsilon)$, or more precisely: $h^{-1}(t; E)$ is defined for $t \in [0, V[E]]$ by

$$h^{-1}(h(\varepsilon; E); E) = \varepsilon , \qquad \text{for } \varepsilon \in [0, \varepsilon^*(E)]$$

$$h^{-1}(t; E) = 0 , \qquad \text{for } t \in [0, V[E \cap S^*]] \ ;$$

$h^{-1}(t)$ is defined for all $t \in [0,\infty)$ by

$$h^{-1}(h(\varepsilon)) = \varepsilon \qquad \text{for} \quad \varepsilon \in [0,\infty)$$

$$h^{-1}(t) \quad = 0 \qquad \text{for} \quad t \in [0, V[E \cap S^*]] \;.$$

Classical properties of inverse functions, and Lemma 2.1 lead to Corollary 2.3.

Corollary 2.3:

Under the same assumptions as in Lemma 2.1, the function $h^{-1}(t; E)$ has the following properties

i) it is continuous for all $t \in [0, V[E]]$ and it is strictly increasing for $t \in [V[E \cap S^*], V[E]]$;

ii) $h^{-1}(t^n; E)$ is convex for all $t \in [0, V^{1/n}[E]]$;

iii) $h^{-1}(t; E) \leq \varepsilon^*(E) \dfrac{t^{1/n} - V^{1/n}[E \cap S^*]}{V^{1/n}[E] - V^{1/n}[E \cap S^*]}$ for all $t \in [V[E \cap S^*], V[E]]$.

The convergence of the ellipsoid method follows from Theorem 2.4, the proof of which is a very minor extension of the proofs given in [2,8,11,22,23].

Theorem 2.4:

Let $\{(x^{(k)}, E_k) : k = 0, 1, \ldots\}$ be a sequence generated by the ellipsoid method applied to a convex function f defined on R^n, which attains its finite minimum f^* on a nonempty set S^*; assume also that $S^* \cap E_0$ is not empty, then:

$$\operatorname*{Min}_{j=0,1,\ldots,k} f(x^{(j)}) \leq f^* + h^{-1}(\tfrac{1}{2} V[E_0] \ (c_n)^k; E_0) \quad .$$

Furthermore there exists a subsequence $\{k_i : i = 0, 1, \ldots\}$ such that:

$$f\left(x^{(k_{i+1})}\right) \leq f\left(x^{(k_i)}\right) , \qquad i = 0, 1, \ldots$$

$$\lim_{i \to \infty} f\left(x^{(k_i)}\right) = f^* ,$$

unless finite termination occurs.

Proof:

Let $\varepsilon_k = h^{-1}(\frac{1}{2} V[E_0] \ c_n^k; E_0)$. We need to show that $f(x^{(j)}) \geq f^* + \varepsilon_k - \delta$ for any $\delta > 0$, and for all $j = 0, 1, \ldots, k$ leads to a contradiction.

Define $D_j = (\bigcap_{i=0}^{j-1} H_i) \cap E_0$, then $D_{j+1} \subset E_j \cap H_j$: clearly this is true for $j = 0$; assume that $D_j \subset E_{j-1} \cap H_{j-1}$, then $D_{j+1} = D_j \cap H_j \subset (E_{j-1} \cap H_{j-1}) \cap H_j \subset E_j \cap H_j$, and the induction is complete.

Also $f(x^{(j)}) \geq f^* + \varepsilon_k - \delta$ for all $j = 0, 1, \ldots, k$ implies that

$$H_j \supset T_{f^*+\varepsilon_k-\delta}$$

$$D_{j+1} \supset (T_{f^*+\varepsilon_k-\delta}) \cap E_0 \quad , \qquad \forall \ j = 0, 1, \ldots, k \ .$$

Hence

$$T_{f^*+\varepsilon_k-\delta} \cap E_0 \subset E_j \cap H_j \qquad \forall \ j = 0, 1, \ldots, k \ ,$$

and thus

$$h(\varepsilon_k - \delta;\ E_0) = V[T_{f^*+\varepsilon_k-\delta} \cap E_0] \leq V[E_k \cap H_k] = \frac{1}{2} V[E_0]\ c_n^k \quad ,$$

or

$$\varepsilon_k - \delta \geq h^{-1}(\frac{1}{2} V[E_0]\ (c_n)^k;\ E_0) \ ,$$

a contradiction.

For the second part of the theorem, define a subsequence $\{k_i : i = 0, 1, \ldots\}$ by induction:

$$k_0 = 0$$

given k_i, define k_{i+1} so that

$$f(x^{(k_{i+1})}) \leq f(x^{(k_i)})$$

and

$$f(x^{(k)}) > f(x^{(k_i)}) \qquad \text{for } k = k_i + 1, \ldots, k_{i+1} - 1;$$

such a k_{i+1} exists by the first part of this theorem (unless $f(x^{(k_i)}) = f^*$).

Q.E.D.

The following theorem indicates that if $V^{1/n}[E_0 \cap S^*] > 0$ (an unusual assumption for an optimization problem), finite convergence occurs, and that if $V^{1/n}[E_0 \cap S^*] = 0$, then convergence is geometric at a rate $c_n^{1/n}$ (which is approximately $1 - 1/2n^2$) which is independent of the function f.

Theorem 2.5:

Under the assumptions of Theorem 2.4, if we let $\varepsilon^*(E_0) = \mathrm{Max}\{f(x) - f^* : x \in E_0\}$, then

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \varepsilon^*(E_0) \frac{(\frac{1}{2} V[E_0](c_n)^k)^{1/n} - V^{1/n}[E_0 \cap S^*]}{V^{1/n}[E_0] - V^{1/n}[E_0 \cap S^*]} ,$$

unless $V[E_0 \cap S^*] > 0$, and

$$k \ge k^* = \left\lceil \log \frac{V[E_0]}{V[E_0 \cap S^*]} \quad \log \frac{1}{c_n} \right\rceil ,$$

in which case $\min_{j=0,\ldots,k} f(x^{(j)}) = f^*$ if $k \ge k^*$; if $V[E_0 \cap S^*] = 0$, then

$$\min_{j=0,1,\ldots k} f(x^{(j)}) \le f^* + (\tfrac{1}{2})^{1/n} \varepsilon^*(E_0)(c_n)^{k/n} .$$

Proof:

This theorem simply combines Theorem 2.4 and Corollary 2.3. Q.E.D.

It should be pointed out that the proof of Theorem 2.4 is a proof based upon a contradiction given by an inclusion of sets (and thus also the volume of sets), but that a number of other characteristics of sets may be used, provided they satisfy a concavity property like that given in Lemma 2.1. For instance the inradius function $r(T_{f^*+\varepsilon} \cap E_0)$ is a concave function of ε, and it would lead to an alternative proof of theorems analogous to Theorems 2.4 and 2.5.

The key condition needed to insure convergence is that $E_0 \cap S^* \ne \phi$. This will fail if:

i) $f^* = -\infty$, and thus $S^* = \phi$;

ii) f^* is finite, and $S^* = \phi$;

iii) f^* is finite, $S^* \neq \phi$, and $S^* \cap E_0 = \phi$.

A derivation similar to the one used earlier gives the following theorem.

Theorem 2.6:

If the ellipsoid method is applied to a general convex function f defined on R^n, and if one does not *assume that* $E_0 \cap S^* \neq \phi$, then

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \leq \mathrm{Min}\{f(x) : x \in E_0\}$$

$$+ [\mathrm{Max}\{f(x) : x \in E_0\} - \mathrm{Min}\{f(x) : x \in E_0\}](c_n)^{k/n} .$$

3. A specific convergence theory

In order to improve the results of Section 2, and to show that the rate of convergence depends upon some characteristics of the function f, it is necessary to impose some conditions on the function f which limit its growth on the whole of R^n. It is known that every convex function is Lipschitz on bounded sets [16, pp. 86 and 237]:

$$|f(x) - f(y)| \leq \Lambda_E \|y-x\| \quad , \qquad \forall\, x,y \in E$$

where

$$\Lambda_E = \mathrm{Sup}\{\|z\| : z \in \partial f(x),\ x \in E\} \quad .$$

Definition 3.1:

A convex function f defined on R^n, which has a nonempty minimum set S^*, is defined to be γ-Lipschitz if there exists a function $\gamma(t)$ defined for all $t \geq 0$, such that $\gamma(0) = 0$, γ is strictly increasing and continuous, and

$$f(x) - f^* \leq \gamma(d(x)) \ , \qquad \text{where } d(x) = d(x,S^*) \ .$$

Clearly, every convex function (with nonempty minimum set S^*) is γ-Lipschitz if γ is defined by:

$$\gamma(d) = \mathrm{Sup}\{f(x)-f^* : d(x,S^*) \leq d\} \quad .$$

Two instances (among many possible ones) of γ functions are given below; they may sometimes be defined a priori for convex functions belonging to a given class.

Condition 3.2:

$$\gamma(d) = \Lambda d^{\alpha} \qquad \text{where} \quad \alpha \geq 1,\ \Lambda > 0 \ .$$

This condition is satisfied with $\alpha = 1$, if:

1) f is piecewise linear, with

$$\Lambda = \mathrm{Sup}\{\|y\| \ : \ y \in \partial f(x),\ x \in R^n\} \quad ;$$

2) f is Lipschitz on the whole of R^n, with

$$\Lambda = \mathrm{Sup}\{\|y\| \ : \ y \in \partial f(x),\ x \in R^n\} \quad .$$

It is satisfied with $\alpha = 2$, if:

1) f is C^2 and there is an upper bound on the largest eigenvalue of the Hessian at any point of R^n;
2) f is C^1 and the gradient of f satisfies a Lipschitz condition on R^n:

$$\|\nabla f(x) - \nabla f(y)\| \leq \frac{\Lambda}{2} \|y-x\| \quad , \qquad \forall\ x,y \in R^n \quad .$$

Condition 3.2 is not satisfied for functions like $f(x_1, x_2) = |x_1| + (x_2)^2$ or $(x_1)^2 + (x_2)^4$; the next condition gives a function γ which has a different degree if d is small, or if d is large.

Condition 3.3:

$$\gamma(d) = \mathrm{Max}(\Lambda_1 d^{\alpha_1}, \Lambda_2 d^{\alpha_2}) \quad ,$$

with $1 \leq \alpha_1 < \alpha_2$ and $\Lambda_1, \Lambda_2 > 0$.

The condition that f is γ-Lipschitz gives an estimate of $\varepsilon^*(E_0) = \mathrm{Max}\{f(x)-f^* : x \in E_0\}$, where $S^* \cap E_0 \neq \phi$ is assumed, which may be used in Theorem 2.5:

$$\varepsilon^*(E_0) = \mathrm{Max}\{f(x)-f^* : x \in E_0\} \leq \mathrm{Max}\{\gamma(d(x)) : x \in E_0\}$$

$$= \gamma(\max_{x \in E_0} d(x)) \leq \gamma(\max_{x,y \in E_0} \|x-y\|) = \gamma(D(E_0)) \ .$$

The theory of convergence of the ellipsoid method on γ-Lipschitz functions will use the properties of the functions:

$$g(\delta; E) = V[(S^* + \delta B) \cap E] \ ,$$

and $g(\delta) = V[S^* + \delta B]$ (defined if S^* is bounded).

If one notices that $(S^* + \delta B) \cap E = \{x \in E : d(x) \leq \delta\}$ and $S^* + \delta B = \{x \in R^n : d(x) \leq \delta\}$, and that d is a convex function with minimum value zero, and minimum set S^*, then it follows that $g(\delta; E)$ and $g(\delta)$ satisfy Lemma 2.1 and Corollary 2.3.

Lemma 3.4:

Let $f(x)$ be a γ-Lipschitz convex function defined on R^n, and E be a compact convex set such that $E \cap S^*$ is not empty, then the volume functions $h(\varepsilon; E)$ and $g(\delta: E)$ satisfy

i) $h(\varepsilon; E) \geq g(\gamma^{-1}(\varepsilon); E), \ \forall\ \varepsilon \geq 0;$

ii) $h^{-1}(t; E) \leq \gamma(g^{-1}(t; E)), \ \ \forall\ t \in [0, V[E]];$

and if S^* is bounded

iii) $h(\varepsilon) \geq g(\gamma^{-1}(\varepsilon)), \ \forall\ \varepsilon \geq 0;$

iv) $h^{-1}(t) \leq \gamma(g^{-1}(t)), \ \forall\ t \geq 0.$

Proof:

Clearly $d(x) \leq \gamma^{-1}(\varepsilon)$ implies

$$f(x)-f^* \leq \gamma(d(x)) \leq \gamma(\gamma^{-1}(\varepsilon)) = \varepsilon \quad ,$$

and thus $S^* + \gamma^{-1}(\varepsilon)B \subset T_{f^*+\varepsilon}$. Taking the volumes of both sides, one gets $g(\gamma^{-1}(\varepsilon)) \leq h(\varepsilon)$. Part i) follows similarly; and parts ii) and iv) are classical properties of inverse functions. Q.E.D.

Theorem 3.5:

Let $f(x)$ be a γ-Lipschitz function defined on R^n; under the same assumptions as in Theorem 2.4, and $V[E_0 \cap S^*] = 0$, the sequence generated by the ellipsoid method satisfies:

$$\min_{j=0,1,\dots k} f(x^{(j)}) \leq f^* + \gamma(D(E_0)\ (\tfrac{1}{2}\ (c_n)^k)^{1/n}) \quad ;$$

if $\gamma = \Lambda d^{\alpha}$ (with $\alpha \geq 1$, $\Lambda > 0$)

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \leq f^* + \Lambda D^{\alpha}(E_0)\,\left(\tfrac{1}{2}\,(c_n)^k\right)^{\alpha/n} ;$$

if $\gamma = \mathrm{Max}(\Lambda_1 d^{\alpha_1}, \Lambda_2 d^{\alpha_2})$ (with $1 \leq \alpha_1 < \alpha_2$; $\Lambda_1, \Lambda_2 > 0$)

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \leq f^* + \mathrm{Max}\{\Lambda_i\, D^{\alpha_i}(E_0)\,\left(\tfrac{1}{2}\,(c_n)^k\right)^{\alpha_i/n} : i = 1,2\} .$$

Proof:

Let $\delta^*(E_0) = \max_{x \in E_0} d(x)$; clearly

$$g(\delta^*(E_0); E_0) = V[E_0] \quad \text{and} \quad g(0, E_0) = V[E_0 \cap S^*] .$$

Lemma 2.3 gives

$$g^{-1}(t; E_0) \leq \delta^*(E_0)\, \frac{t^{1/n} - V^{1/n}[E_0 \cap S^*]}{V^{1/n}[E_0] - V^{1/n}[E_0 \cap S^*]} .$$

But $\delta^*(E_0) \leq D(E_0)$ and thus

$$h^{-1}(t; E_0) \leq \gamma(g^{-1}(t); E_0) \leq \gamma\left(D(E_0)\, \frac{t^{1/n} - V^{1/n}[E_0 \cap S^*]}{V^{1/n}[E_0] - V^{1/n}[E_0 \cap S^*]}\right) ,$$

and the theorem follows from Theorem 2.4. Q.E.D.

If $V[E_0 \cap S^*] \neq o$, then a finite convergence result identical to Theorem 2.5 is of course true.

If $\alpha = 1$, this theorem gives the same result as Theorem 2.5; but if $\alpha > 1$, this theorem gives a rate of convergence of $(c_n)^{\alpha/n}$ which is α times faster than the rate given by Theorem 2.5.

The convergence results given above are all based upon the concavity of $h^{1/n}$ or $g^{1/n}$; this implied that

$$h(\varepsilon;\ E) \geq [\frac{\varepsilon}{\varepsilon^*(E)} V^{1/n}[E] + (1 - \frac{\varepsilon}{\varepsilon^*(E)}) V^{1/n}[E \cap S^*]]^n$$

(and something similar for g). The right hand side is clearly a polynomial of degree n in ε, with all of its $n+1$ coefficients positive if $0 < V[E \cap S^*] < V[E]$, while if $V[E \cap S^*] = 0$ all coefficients but that of ε^n are zero. Some sharper results exist for the functions g, and if the function f is γ-Lipschitz this imparts related properties on h.

The function $g(\delta) = V(S^* + \delta B]$ has been studied extensively (if S^* is compact and convex); the key result due to Steiner and Minkowsky (see Theorem 4.9) is that $g(\delta)$ is a polynomial of degree n in δ, with nonnegative coefficients and the coefficients of ε^j $(j = 0, 1, \ldots, n-s-1$, with $s = \dim S^*)$ are all zero.

The function $g(\delta;\ E) = V[(S^* + \delta B) \cap E]$ has not received much attention, and is much harder to study; Theorem 4.10 provides a bound on $g(\delta;\ E)$ which is a polynomial, with some negative coefficients which still permits an improvement of the convergence theory given above.

The main results given below can be summarized by saying that the rate of convergence of the ellipsoid method, if f is γ-Lipschitz (and

$\gamma = \Lambda d^{\alpha}$) and the dimension of $S^* \cap E$ is s, is equal to $(c_n)^{\alpha/(n-s)}$.

The analysis is not done if $s = 0$ or $s = n$, as it can be seen that no significant improvement on Theorems 2.5 and 3.5 occurs; by this is meant that the rate of convergence would not be improved. The only exception would be the unlikely case where $\dim(S^* \cap E_0) = 0$, and $\dim S^* \geq 1$.

Theorem 3.6:

Let f be a γ-Lipschitz function defined on R^n; under the same assumptions as in Theorem 2.4, plus the fact that $E_0 \cap S^*$ is of dimension $s \geq 1$, $s \leq n-1$, the sequence generated by the ellipsoid method satisfies:

1) If $s = 2, \ldots, n-1$

$$\underset{j=0,1,\ldots,k}{\text{Min}} f(x^{(j)}) \leq f^* + \gamma\left(\left[\frac{(n-s+1)\ V[E_0]}{2\omega_{n-s}\ V_s[E_0 \cap S^*]}\ (c_n)^k\right]^{1/(n-s)}\right)$$

and if $\gamma = \Lambda d^{\alpha}$

$$\underset{j=0,1,\ldots,k}{\text{Min}} f(x^{(j)}) \leq f^* + \Lambda\left[\frac{(n-s+1)\ V[E_0]}{2\omega_{n-s}\ V_s[E_0 \cap S^*]}\right]^{\alpha/n-s} (c_n^{\alpha/(n-s)})^k$$

provided that k is large enough:

$$\frac{1}{2}\ V[E_0]\ (c_n)^k \leq \frac{\omega_{n-s}\ V_s[E_0 \cap S^*]}{n-s+1}\left[\frac{n-s}{n-s+1}\ \frac{V_s[E_0 \cap S^*]\ r(E_0)}{s\omega_s\left(\frac{D(E_0)}{2}\right)^s}\right]^{n-s} .$$

2) If $s = 1$

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \gamma\left(\left[\frac{1}{2}\frac{(2n-1)\ V[E_0]}{\omega_{n-1}\ V_1[E_0 \cap S^*]}\ (c_n)^k\right]^{1/(n-1)}\right)$$

and if $\gamma = \Lambda d^{\alpha}$

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \Lambda\left[\frac{1}{2}\frac{(2n-1)\ V[E_0]}{\omega_{n-1}\ V_1[E_0 \cap S^*]}\right]^{\alpha/(n-1)} (c_n^{\alpha/(n-1)})^k$$

provided that k is large enough:

$$\frac{1}{2}\ V[E_0]\ (c_n)^k \le \frac{\omega_{n-1}\ V[E_0 \cap S^*]}{2n-1}\left[\frac{(n-1)}{\sqrt{2}\ (n-\frac{1}{2})}\ \frac{V_1[E_0 \cap S^*]\ \sqrt{r(E_0)}}{D(E_0)}\right]^{n-1} .$$

<u>Proof</u>:

The analysis is given only for $s = 2, \ldots, n-1$ (it is analogous if $s = 1$, and useless if $s = n$).

Theorem 4.10 implies that

$$g(\delta;\ E_0) \ge \omega_{n-s}\left[V_s[E_0 \cap S^*] - s\left(\frac{D(E_0)}{2}\right)^s \omega_s\ \frac{\delta}{r(E_0)}\right] \delta^{n-s} \quad \forall\ \delta \in [0,\ r(E_0)] .$$

Now if η, k, $\ell > 0$, it is easy to check that

$$\delta^k - \eta\delta^{k+\ell} \ge \frac{\ell}{k}\ \delta^k \qquad \text{for all } \delta \in \left[0,\ \left(\frac{1}{\eta}\ \frac{k}{k+\ell}\right)^{1/\ell}\right] .$$

Thus

$$g(\delta;\ E_0) \geq \omega_{n-s} \frac{V_s[E_0 \cap S^*]}{n-s+1} \delta^{n-s}$$

for all $\delta \in [0,\ \tilde{\delta}(E_0)]$ where

$$\tilde{\delta}(E_0) = \frac{n-s}{n-s+1} \frac{V_s[E_0 \cap S^*]\ r(E_0)}{s\omega_s(\frac{D(E_0)}{2})^s} .$$

and

$$g^{-1}(t;\ E_0) \leq \left[\frac{(n-s+1)t}{\omega_{n-s}\ V_s[E_0 \cap S^*]}\right]^{1/(n-s)}$$

for all $t \in [0,\ t^*(E_0)]$ where

$$t^*(E_0) = \omega_{n-s} \frac{V_s[E_0 \cap S^*]}{n-s+1} \tilde{\delta}(E_0))^{n-s} .$$

The theorem follows from Lemma 3.4 and Theorem 2.4. Q.E.D.

The next convergence results are based upon the Steiner polynomial $g(\delta)$ (see Section 4). The order to use $g(\delta)$ *is the study of the* convergence of the ellipsoid method one must assume that S^* is compact and that E_0 contains entirely some level sets of f: define

$$\varepsilon_0 = \text{Sup}\{\varepsilon : T_{f^*+\varepsilon} \subset E_0\} = \text{Inf}\{f(x)-f^* : x \notin E_0\} .$$

If f is γ-Lipschitz then

$$S^* + \gamma^{-1}(\varepsilon)B \subset E_0 \qquad \forall\, \varepsilon \in [0,\, \varepsilon_0]$$

and

$$g(\delta;\, E_0) = g(\delta) = \sum_{i=0}^{n} \binom{n}{i} W_i[S^*]\, \delta^i \qquad \forall\, \delta \in [0,\, \gamma^{-1}(\varepsilon_0)] \quad .$$

If $V[S^*] < \frac{1}{2} V[E_0]\, (c_n)^k$, define $\delta(k)$ to be the unique positive zero of

$$\sum_{i=0}^{n} \binom{n}{i} W_i[S^*]\, \delta^i - \frac{1}{2} V[E_0]\, (c_n)^k \quad ,$$

then $\min\limits_{j=0,\ldots,k} f(x^{(j)}) \leq f^* + \gamma(\delta(k))$ provided that $\delta(k) \leq \gamma^{-1}(\varepsilon_0)$.

If $V[S^*] \geq \frac{1}{2} V[E_0]\, (c_n)^k$, then there are no positive zeros, implying that finite convergence has occurred.

Clearly, as $g(\delta)$ is a polynomial with nonnegative coefficients, one has:

$$\delta(k) \leq \left(\frac{V[E_0]\, (c_n)^k}{2 \binom{n}{i} W_i[S^*]} \right)^{1/(n-i)}$$

if $i = n-s, \ldots, n$ and $i \neq 0$, where $s = \dim S^*$.

Theorem 3.7.

Let f be a γ-Lipschitz function defined on R^n; under the same assumptions as in Theorem 2.4 plus the facts that S^* is compact, and that $\varepsilon_0 = \mathrm{Inf}\{f(x)-f^* : x \notin E_0\}$ is positive, then the sequence generated by the ellipsoid method satisfies:

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \min_{\substack{n-s \le i \le n \\ i \ne 0}} \gamma\left(\left[\frac{V[E_0]\ (c_n)^k}{2\binom{n}{i}\ W_i[S^*]}\right]^{1/(n-i)}\right)$$

where $s = \dim S^*$, provided that

$$\min_{\substack{n-s \le i \le n \\ i \ne 0}} \gamma\left(\left[\frac{V[E_0]\ (c_n)^k}{2\binom{n}{i}\ W_i[S^*]}\right]^{1/(n-i)}\right) \le \varepsilon_0 \ ;$$

if in addition $\gamma = \Lambda d^\alpha$ and $s \le n-1$, then:

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \Lambda\left[\frac{V[E_0]}{2\omega_{n-s}\ V_s[S^*]}\right]^{\alpha/(n-s)} (c_n^{\alpha/(n-s)})^k$$

provided that

$$\Lambda\left[\frac{V[E_0]\ (c_n)^k}{2\omega_{n-s}\ V_s[S^*]}\right]^{\alpha/(n-s)} \le \varepsilon_0 \quad .$$

Thus every nonzero coefficient of the Steiner polynomial gives a bound on the convergence of the ellipsoid method; for instance if $\gamma = \Lambda d^\alpha$:

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \Lambda\left[\frac{V[E_0]\ (c_n)^k}{\omega_{n-s+1}\ F_s[S^*]}\right]^{\alpha/(n-s+1)} , \quad \text{if} \quad s \ge 2 \ ;$$

or

$$\min_{j=0,1,\ldots,k} f(x^{(j)}) \le f^* + \Lambda\left[\frac{V[E_0]\ (c_n)^k}{2\omega_n}\right]^{\alpha/n} , \qquad \text{if} \quad s = n$$

provided that k is large enough.

It might be worth pointing out that every result in this paper (except for Theorem 3.6) is valid for Levin's method of centers of gravity, if c_n is replaced by $\tilde{c}_n$.

4. Volumes of convex arrays

Definition 4.1:

A mapping T_u from a convex subset U of a linear space onto the class of convex subsets of R_n is called a concave array if

$$u_1, u_2 \in U \qquad \lambda \in [0,1]$$

implies

$$T_{\lambda u_1 + (1-\lambda) u_2} \supset \lambda T_{u_1} + (1-\lambda)\, T_{u_2} \quad ;$$

it is a convex array if the inclusion is reversed, and a linear array if equality holds.

Clearly a linear array is both convex and concave.

Lemma 4.2:

Let A and A_i, $i = 1, \ldots, r$, be convex subsets of R^n, and E be a convex subset of R^n, then

1) $\left(\sum_{i=1}^{r} \nu_i\right) A = \sum_{i=1}^{r} \nu_i A$ if $\nu = (\nu_1, \ldots, \nu_r) \in R_+^r$ (or $\nu \in -R_+^r$).

2) $\sum_{i=1}^{r} \nu_i (A_i \cap E) \subset \left(\sum_{i=1}^{r} \nu_i A_i\right) \cap E$ if $\sum_{i=1}^{r} \nu_i = 1$ and $\nu \in R_+^r$.

3) $\sum_{i=1}^{r} \nu_i A_i$ is a convex array for $\nu \in R^r$; it is a linear array if $\nu \in R_+^r$ (or any of the 2^r orthants of R^r).

Lemma 4.3:

Let T_u be a concave (or linear) array defined on a convex set U, and E be a convex subset of R^n, then $T_u \cap E$ is a concave array on U.

Lemma 4.4:

Let f be a convex function defined on R^n, then $T_\alpha = \{x \in R^n : f(x) \leq \alpha\}$ is a concave array for $\alpha \in R$; and if E is a convex set $T_\alpha \cap E$ is also a concave array.

Lemma 4.5:

Let E and A_i, $i = 1, \ldots, r$ be convex subsets of R^n, then $(\sum_{i=1}^{r} \nu_i A_i) \cap E$ is a concave array on $\nu = (\nu_1, \ldots, \nu_r) \in R_+^r$ (or any of the 2^r orthants of R^r).

Theorem 4.6:

Let f be a convex function defined on R^n, E a compact convex subset of R^n, $T_\alpha = \{x \in R^n : f(x) \leq \alpha\}$, and $\alpha^* = \mathrm{Min}\{f(x) : x \in E\}$; then $T_\alpha \cap E$ is continuous in α for $\alpha \in [\alpha^*, \infty)$.

Proof:

As f is continous, $T_\alpha \cap E$ is convex, compact and nonempty for all $\alpha \geq \alpha^*$ (and empty for $\alpha < \alpha^*$).

First we show that for all $\alpha \geq \alpha^*$

$$\lim_{\eta \downarrow 0} T_{\alpha+\eta} \cap E = T_\alpha \cap E \; ; \qquad \text{i.e., for each } \delta > 0$$

there exists an $\tilde{\eta} > 0$ such that for all $\eta \in (0, \tilde{\eta}]$, $(T_{\alpha+\eta} \cap E) + \delta B^0 \supset T_\alpha \cap E$ and $(T_\alpha \cap E) + \delta B^0 \supset T_\alpha$. The first inclusion follows from $T_{\alpha+\eta} \supset T_\alpha$ for $\eta \geq 0$. The second inclusion is not true if and only if $K_\eta = ((T_\alpha \cap E) + \delta B^0)^c \cap (T_{\alpha+\eta} \cap E)$ is not empty for all $\eta > 0$ (A^c denotes

the complement of A in R^n). The sets K_η for $\eta > 0$ form a nested sequence of compact sets, and thus there exists an $x \in R^n$ such that $x \in K_\eta$, $\eta > 0$ (unless some K_η as empty). Such an x belongs to $T_{\alpha+\eta} \cap E$ for all $\eta > 0$, and thus $x \in E$ and $x \in T_\alpha$; also $x \in (T_\alpha \cap E) + \delta B^{0})^c$, and thus $x \notin T_\alpha \cap E$, a contradiction, concluding the first part of this proof. Note that no convexity assumptions are needed for E or f.

For the second part (i.e., $\lim_{\eta \downarrow 0} T_{\alpha-\eta} \cap E = T_\alpha \cap E$ for $\alpha > \alpha^*$ and $\eta \leq \alpha-\alpha^*$) one needs to show that for every $\delta > 0$, there exists an $\tilde{\eta} > 0$ ($\tilde{\eta} \leq \alpha-\alpha^*$) such that $(T_\alpha \cap E) + \delta B^0 \supset T_{\alpha-\eta} \cap E$ and $(T_{\alpha-\eta} \cap E) + \delta B^0 \supset T_\alpha \cap E$ for all $\eta \in (0,\tilde{\eta}]$. The first inclusion is trivial; if the second is not true, then, following the reasoning used above, there exists an $x \in R^n$ such that

$$x \in ((T_{\alpha-\eta} \cap E) + \delta B^0)^c \cap (T_\alpha \cap E) \qquad \text{for } \eta \in (0,\tilde{\eta}] ;$$

i.e., $x \in T_\alpha \cap E$, $x \notin (T_{\alpha-\eta} \cap E) + \delta B^0$ for $\eta \in (0,\tilde{\eta}]$.

Choose $\tilde{x} \in T_{\alpha-\tilde{\eta}} \cap E$; it follows from the convexity of E and f that $\lambda\tilde{x} + (1-\lambda)x \in T_{\alpha-\lambda\tilde{\eta}} \cap E$ if $\lambda \in [0,1]$; take $\tilde{\lambda}$ such that $0 < \tilde{\lambda} < \delta/\|\tilde{x}-x\|$, then $x^* = \tilde{\lambda}\tilde{x} + (1-\tilde{\lambda})x$ satisfies:

$$\|x^*-x\| < \delta \quad \text{and} \quad x^* \in T_{\alpha-\lambda\tilde{\eta}} \cap E .$$

Thus $x = x^* + (x-x^*) \in (T_{\alpha-\lambda\tilde{\eta}} \cap E) + \delta B^0$, a contradiction. Q.E.D.

Theorem 4.7 (the Brunn-Minkowski inequality):

Let $T_u \in U$ (where U is a convex set) be a concave array of compact convex subsets of R^n, then $V^{1/n}[T_u]$ is a concave function of u.

Proof: [3], p. 88. [5], p. 97, [9], pp. 159 and 187.

Theorem 4.8:

Let A_i, $i = 1, \ldots, r$, be r compact convex subsets of R^n, and ν_i, $i = 1, \ldots, r$ be r nonnegative numbers, then the volume of $\sum_{i=1}^{r} \nu_i A_i$ is an homogeneous n^{th} degree polynomial in the variables $\nu_1, \ldots, \nu_r$,

$$V[\sum_{i=1}^{r} \nu_i A_i] = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_r=1}^{n} V[A_{i_1}, A_{i_2}, \ldots, A_{i_n}] \nu_{i_1} \nu_{i_2} \cdots \nu_{i_n}$$

where the coefficients $V[A_{i_1}, A_{i_2}, \ldots, A_{i_n}]$ are chosen to be invariant under permutations of the arguments, and are called the mixed volumes of the linear array. The mixed volumes are nonnegative, and continuous and increasing functions of their arguments.

Proof: [3], p. 40; [5], p. 85.

Of particular use will be the linear array of parallel sets $A+\delta B$, $\delta \geq 0$, with A a compact convex set, and B the closed unit ball in R^n. The mixed volume $V[\underbrace{A, \ldots, A}_{n-k}, \underbrace{B, \ldots, B}_{k}]$ is called the k^{th} transversal

mass integral and denoted as $W_k[A]$. It follows immediately from Theorem 4.8, that

$$V[A+\delta B] = \sum_{k=0}^{n} \binom{n}{k} W_k[A]\ \delta^k$$

is a n^{th} degree polynomial in δ with positive coefficients. It can be shown that $W_0[A] = V[A]$, $nW_1[A] = F[A]$ the surface of A, and $W_n[A] = \omega_n$.

Also $[W_k[T_u]]^{1/(n-k)}$ is a concave function of u, if T_u, $u \in U$ is a concave array [1, 6, 7].

If A is a s dimensional compact convex set on R^n then the following theorem characterizes $V[A+\delta B]$.

<u>Theorem 4.9</u>:

Let A be a s dimensional compact convex subset of R^n, and let $\tilde{W}_k[A]$, $k = 0, 1, \ldots, s$ be the transversal mass integrals of A seen as a subset of a s dimensional space, then:

$$V[A+\delta B] = \sum_{k=n-s}^{n} \frac{\omega_k}{\omega_{k-n+s}} \binom{s}{k-n+s} \tilde{W}_{k-n+s}[A]\ \delta^k \quad ;$$

the coefficients of δ^{n-s}, δ^{n-s+1} and δ^n are respectively $\omega_{n-s} V_s[A]$ $\frac{1}{2}\,\omega_{n-s+1}\, F_s(A)$ and ω_n.

Proof:

By induction on [9] pp. 215-216, one has

$$W_k[A] = \frac{\kappa_k}{\kappa_{k-n+s}} \frac{\binom{s}{k-n+s}}{\binom{n}{k}} W_{k-n+s}[A]$$

for $k = n-s, \ldots, n$; and $W_k[A] = 0$ for $k = 0, \ldots, n-s-1$. Clearly $W_0[A] = V_s[A]$, $sW_1[A] = F_s[A]$ and $W_s[A] = \kappa_s$. Q.E.D.

Theorem 4.10:

Let A be a closed convex subset of R^n, E an ellipsoid in R^n, and $A \cap E$ be of dimension $s \geq 1$ (and thus A is of dimension s), then

i) if $s \geq 2$

$$V[(A+\delta B) \cap E] \geq \kappa_{n-s}[V_s[A \cap E] - s(\frac{D(E)}{s})^s \kappa_s \frac{\delta}{r(E)}] \delta^{n-s};$$

ii) if $s = 1$

$$V[(A+\delta B) \cap E] \geq \kappa_{n-1}[V_1[A \cap E] - D(E) \sqrt{\frac{2\delta}{r(E)}}] \delta^{n-1}$$

where $0 \leq \delta \leq r(E)$.

Proof:

Assume, without restriction, that E has its center at the origin of coordinates.

Define $A_\eta = A \cap (1 - \frac{\eta}{r(E)})E$ for $\eta \in [0, r(E)]$; clearly $A_\eta + \delta B \subset E$ for $\delta \in [0,\eta]$. Thus

$$V[(A_\eta + \delta B) \cap E] = V[A_\eta + \delta B] \geq \omega_{n-s} V_s[A_\eta] \delta^{n-s} \quad \forall \quad 0 \leq \delta \leq \eta \leq r(E)$$

(using Theorem 4.9). Now

$$V_s[A \cap E] - V_s[A_\eta] = V_s[A \cap (E/(1 - \frac{\eta}{r(E)}) E]$$

$$\leq V_s[M(A) \cap (E/(1 - \frac{\eta}{r(E)}) E] \quad \text{if } \eta \in [0, r(E)],$$

where $M(A)$ denotes the s-dimensional affine manifold containing A, and $/$ denotes the set difference $(S_1/S_2 = \{x \in S_1 : x \notin S_2\})$.

Define a linear map T such that $T(E) = B$, and let $L(A) = M(A) - A$ be the s-dimensional linear subspace parallel to $M(A)$, then:

$$V_s[M(A) \cap (E/(1 - \frac{\eta}{r(E)}) E)]$$

$$= \frac{V_s[L(A) \cap E]}{V_s[T(L(A) \cap E)]} V_s[T(M(A) \cap (E/(1 - \frac{\eta}{r(E)}) E))]$$

$$= \frac{V_s[L(A) \cap E]}{\omega_s} V_s[M(T(A)) \cap (B/(1 - \frac{\eta}{r(E)}) B)]$$

$$\leq (\frac{D(E)}{2})^s V_s[M(T(A)) \cap (B/(1 - \frac{\eta}{r(E)}) B)]$$

where $\eta \in [0, r(E)]$.

Now let $t = d(0, M(T(A)))$, where clearly $t \in [0,1]$ as $A \cap E$ is not empty; then

$$V_s[M(T(A)) \cap (B/(1 - \frac{\varepsilon}{r(E)})\, B]$$

$$= \begin{cases} \omega_s\left[(1-t^2)^{s/2} - (1 - \frac{\varepsilon}{r(E)})^s \left(1 - \left(\frac{t}{1 - \frac{\varepsilon}{r(E)}}\right)^2\right)^{s/2}\right] & \text{if } 0 \le t \le 1 - \varepsilon/r(E) \\ \omega_s(1-t^2)^{s/2} & \text{if } 1 - \varepsilon/r(E) \le t \le 1 . \end{cases}$$

The maximum of this as a function of t is attained for $t = 0$ if $s > 2$, and for $t = 1 - \varepsilon/r(E)$ if $s = 1$ or 2 (note: if $s = 2$ then it is constant for $t \in [0, 1 - \varepsilon/r(E)]$). *Thus*

$$V_s[M(T(A)) \cap (B/(1 - \frac{\varepsilon}{r(E)})\, B)]$$

$$\le \begin{cases} \omega_s[1 - (1 - \varepsilon/r(E))^s] , & \text{if } s \ge 2 \\ \omega_1[1 - (1 - \varepsilon/r(E))^2]^{1/2} , & \text{if } s = 1 \end{cases}$$

$$\le \begin{cases} s\,\omega_s \frac{\varepsilon}{r(E)} , & \text{if } s \ge 2 \\ \omega_1\left(\frac{2\varepsilon}{r(E)}\right)^{1/2} , & \text{if } s = 1 \end{cases}$$

It follows that

$$V_s[A_\delta] \leq V_s[A \cap E] - \begin{cases} \left(\frac{D(E)}{2}\right)^s \; s\omega_s \frac{\delta}{r(E)} \;, & \text{if } s \geq 2 \\ D(E) \sqrt{\frac{2\eta}{r(E)}} \;, & \text{if } s = 1 \;, \end{cases}$$

and the theorem follows. Q.E.D.

If $s = \dim(A \cap E) = 0$, and A reduces to a point in the interior of E, then $V[(A+\delta B) \cap E] = \omega_n \delta^n$ for δ small enough; if $s = \dim(A \cap E) = 0$, then A itself may be of any dimension, and a slightly different approach would be required.

5. Acknowledgments

We wish to thank Richard E. Stone for pointing to us the existence of the translation of the two papers by Yudin and Nemirovskii.

REFERENCES

1. Alexandrow, A.D., "Neue Ungleichungen fur die Mischvolumen Konvexe Korper," C.r. Acad. Sci. USSR, n. Ser 14, 155-157 (1937).

2. Aspvall, B. and Stone, R.E., "Khachiayan's Linear Programming Algorithm," Journal of Algorithms, 1, No. 1 (1980) 1-13.

3. Bonnensen, T. and Fenchel, W., "Theorie der Konvexen Korper," Springer, Berlin, 1934. Reprint: Chelsea, New York 1948.

4. Busemann, H., "Convex Surfaces," Interscience Publishers Inc., New York-London 1958.

5. Eggleston, H.G., "Convexity," Cambridge University Press 1958.

6. Fenchel, W., "Inegalites Quadratiques Entre Les Volumes Mixtes des Corps Convexes," Comptes Rendus de L'Academie des Sciences (Paris) 203 (1936) 647-650.

7. Fenchel, W., "Generalisation du Theoreme de Brunn et Minkowski Concernant les Corps Convexes," Comptes Rendus de L'Academie des Sciences (Paris) 203 (1936) 764-766.

8. Gacs, P. and Lovasz, L., "Khachian's Algorithm for Linear Programming," Report CS-79-750, Computer Science Dept., Stanford University, Stanford, CA, U.S.A. (1979).

9. Hadwiger, H., "Vorlesungen Uber Inhalt, Oberflache und Isoperimetrie," Springer Verlag, Berlin-Gottingen-Heidelberg 1957.

10. Khacian, L.G., "A Polynomial Algorithm in Linear Programming," Doklady Akademiia Nauk USSR, 244, No. 5 (1979) 1093-1095. Translated in Soviet Mathematics Doklady, 20, No. 1 (1979) 191-194.

11. Khacian, L.G., "A Polynomial Algorithm in Linear Programming," Z. Vycisl. Mat. 1 Mat. Fiz., 20 No. 1 (1980) 51-68 (in Russian).

12. Korte, B. and Schrader, R., "A Note on Convergence Proofs for Shor-Khacian Methods," Report 80156-OR, Institut fur Operations Research, Universitat Bonn, Bonn, W. Germany.

13. Levin, A.Y., "On an Algorithm for the Minimization of Convex Functions," Doklady Akademiia Nauk USSR, 160, No. 6 (1965) 1244-1247. Translated as Soviet Mathematics Doklady 6, No. 1 (1965) 286-290.

14. Mityagin, B.S., "Two Inequalities for Volumes of Convex Bodies," Matematicheskie Zametki, 5, No. 1 (1969) 99-106. Translated as Mathematical Notes of the Academy of Sciences of the USSR 5, No. 1 (1969) 61-65.

15. Nemirovskii, A.S. and Yudin, D.B., "Optimization Methods Adapting to the Problem of 'Significant' Dimension," Avtomatika i Telemekhanika 38, No. 4, 75-87, 1977. Translated in Automation and Remote Control 38, No. 4, 513-524, 1977.

16. Rockafellar, R.T., "Convex Analysis," Princeton University Press 1970.

17. Shor, N.Z., "Utilization of the Operation of Space Dilatation in the Minimization of Convex Functions," Kibernetika 6 (1) (1970) 6-12. Transated as Cybernetcis 6 (1) (1970), 7-15.

18. Shor, N.Z., "Generalized Gradient Methods for Non-smooth Functions and their Applications to Mathematical Programming Problems," Ekonomika i Matematicheskie Metody, 12 (2) (1976) 337-356 (in Russian).

19. Shor. N.Z., "New Development trends in Nondifferentiable Optimization," Kibernetika 13 (6) (1977) 87-91. Translated as Cybernetics 13 (6) 881-886.

20. Shor, N.Z., "Cut-off methods with space extension in convex programming problems," Kibernetika 13 (1) (1977) 94-95. Translated as Cybernetics 13 (1) (1977) 94-96.

21. Shor, N.Z. and Gershovich, V., "On one Family of Solution Algorithms for Problems of Convex Programming," Kibernetika 15 (4) (1979), 62-67. Translated as Cybernetics 15 (4) (1979) 502-508.

22. Yudin, D.B. and Nemirovskii, A.S., "Estimation of the Informational Complexity of Mathematidal Programming Problems," Economika i Matematicheskie Metody 12, No. 1 (1976) 128-142. Translated in Matekon 13, No. 2 (1977) 3-25.

23. Yudin, D.B. and Nemirovskii, A.S., "Informational Complexity and Effective Methods for the Solution of Convex Extremal Problems," Ekonomika i Matematicheskie Metody 12, No. 2 (1976) 357-369. Translated in Matekon 13, No. 3 (1977) 25-45.

24. Yudin, D.B. and Nemirovskii, A.S., "Informational Complexity for Strict Convex Programming," Ekonomika i Matematicheskie Metody 13 (3) (1977) 550-559 (in Russian).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 SOL-81-3	2. GOVT ACCESSION NO.: AD-A099 292
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): CONVERGENCE RATES OF THE ELLIPSOID METHOD ON GENERAL CONVEX FUNCTIONS	5. TYPE OF REPORT & PERIOD COVERED: 9 Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 Jean-Louis Goffin 12	8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0267; DE-[illegible]-76SF [illegible]
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Operations Research - SOL, Stanford University, Stanford, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-047-143
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research - Dept. of the Navy 11, 800 N. Quincy Street, Arlington, VA 22217	12. REPORT DATE: January 1981
	13. NUMBER OF PAGES: 38 12 4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ellipsoid method
nondifferentiable optimization
convex programming
volumes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

The ellipsoid method is applied to the unconstrained minimization of a general convex function. The method converges at a geometric rate, which depends only upon the dimension of the space but not on the actual function. This rate can be improved somewhat if the function satisfies some Lipschitz-type condition, or if the minimum set has dimension greater than zero.

If the ellipsoid entirely contains the optimal set, equating the Steiner polynomial associated to the optimal set, and the volume of the ellipsoid at a given iteration, will give an upper bound on the minimum recorded function value.

FILMED

6-8